DOCUMENT RESUME

ED 451 408 CE 081 583

AUTHOR Wagner, Daniel A.
TITLE Literacy Assessment for Out-of-School Youth and Adults: Concepts, Methods and New Directions. ILI Technical Report.
INSTITUTION International Literacy Inst., Philadelphia, PA.
REPORT NO ILI-TR98-09
PUB DATE 1998-09-00
NOTE 23p.; Paper was originally prepared for the Expert Seminar on Literacy in Out-of-School Youth and Adults (Paris, France, June 22-24, 1998), co-sponsored by the International Literacy Institute (ILI) and UNESCO.
AVAILABLE FROM For full text: http://literacyonline.org/products/ili/pdf/TR9809.pdf.
PUB TYPE Opinion Papers (120) -- Reports - Evaluative (142)
EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Academic Achievement; Adolescents; Adult Basic Education; *Adult Literacy; Comparative Analysis; *Comparative Education; Data Collection; Developing Nations; *Educational Assessment; *Evaluation Methods; Foreign Countries; Global Approach; *International Cooperation; National Surveys; Out of School Youth; Secondary Education; *Skill Development; Vocabulary

ABSTRACT

A critical need remains for a practical and low-cost methodology to bridge the gap between methodologies for assessment of literacy at the national survey level. At the program level, it must be able to be used effectively by those in developing countries with limited funds but major literacy problems. Etic (external, quantifiable, comparison-oriented) measures are clearly important in understanding how people acquire literacy, how educators and policymakers view literacy, and how economic and societal systems interact in an increasingly interdependent world. Collection of international statistics on literacy poses problems that specialists have debated. With the greater need for direct measurement and increased technical capacity for such measurement in developing countries, clarification of issues is occurring in language policy and multilingualism; classification; international comparability of data; and measuring the consequences of literacy. The Education for All approach that emphasizes measuring learning achievement and widening the view to beyond the 3Rs raises definition issues. National and local program needs are artificially separated by goals and tools typically used in these two domains. Confusion exists with regard to the terminology used to describe methods for literacy assessment. Challenges to validity of literacy assessment tools are cross-cultural comparison, targets and "minimum" levels of achievement, and technical/statistical challenges. (Contains 51 references.) (YLB)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 451 408

TECHNICAL REPORT

Literacy Assessment for Out-of-School Youth and Adults: Concepts, Methods, and New Directions

IEI Technical Report IR98-09
(September 1998)

International Literacy Institute
Graduate School of Education
University of Pennsylvania

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY:

J. Smith

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

1

BEST COPY AVAILABLE

CE 061 583

INTERNATIONAL LITERACY INSTITUTE

The International Literacy Institute (ILI), officially co-sponsored by the United Nations Educational, Scientific, and Cultural Organization (UNESCO) and the University of Pennsylvania Graduate School of Education, was formally established in 1994 in Philadelphia. The ILI builds on more than a dozen years of university-based literacy research, the federally funded National Center on Adult Literacy (NCAL), and close collaboration with governmental, non-governmental, and multilateral agencies worldwide. The ILI and NCAL share the same building with the Penn Technology in Education Learning Laboratory (PennTELL) on the campus of the University of Pennsylvania.

The ILI mission is to provide scientific leadership in training and development in literacy, with a special emphasis on developing countries. The ILI defines literacy as primarily the basic skills of reading, writing, and mathematics among children, youth, and adults. The term also includes lifelong and work-related skills.

The ILI has had extensive experience in evaluation and applied research on basic education and literacy in developing countries around the world. The ILI's development activities include partnerships to foster regional and national centers of excellence; training to enhance the capacity of national and regional institutions and agencies; innovations derived from research, development, and evaluation; information dissemination that provides a forum for the exchange of ideas; advanced technologies to increase communication and learning achievement; and training and development activities in both formal school settings and nonformal education programs.

UNITED NATIONS EDUCATIONAL, SCIENTIFIC AND CULTURAL ORGANIZATION

The constitution of the United Nations Educational, Scientific and Cultural Organization (UNESCO) was adopted by the London Conference in November 1945, and entered into effect on 4 November 1946 when 20 states had deposited instruments of acceptance. It currently has 186 Member States.

The main objective of UNESCO is to contribute to peace and security in the world by promoting collaboration among nations through education, science, culture, and communication in order to further universal respect for justice, for the rule of law, and for the human rights and fundamental freedoms which are affirmed for the peoples of the world, without distinction of race, sex, language, or religion, by the Charter of the United Nations.

To fulfill its mandate, UNESCO performs five principal functions:

- *Prospective studies*: what forms of education, science, culture, and communication for tomorrow's world?
- *The advancement, transfer, and sharing of knowledge*: relying primarily on research, training, and teaching activities.
- *Standard-setting action*: the preparation and adoption of international instruments and statutory recommendations.
- *Expertise*: provided to Member States for their development policies and projects in the form of "technical co-operation."
- *Exchange of specialized information*.

(This description is adapted from text on the UNESCO website, found at www.unesco.org.)

The ILI receives major funding support from a number of agencies and organizations, including UNESCO, the Ministry of Foreign Affairs of Norway, and the University of Pennsylvania. Neither these agencies nor the ILI may be held responsible for the particular views of the individual authors of the ILI reports.

© Copyright 1998 International Literacy Institute

LITERACY ASSESSMENT FOR OUT-OF-SCHOOL YOUTH AND ADULTS:

Concepts, Methods, and New Directions

Daniel A. Wagner
International Literacy Institute
University of Pennsylvania

Author Biography

Daniel A. Wagner is Professor of Education at the University of Pennsylvania and Director of the International Literacy Institute (ILI). He is also Director of the National Center on Adult Literacy at the University of Pennsylvania. He received his Ph.D. in educational psychology at the University of Michigan, was a two-year postdoctoral fellow at Harvard University, a Visiting Fellow at the International Institute of Education Planning in Paris, and a Visiting Professor at the University of Geneva (Switzerland).

Prefatory note

This paper was originally prepared for the Expert Seminar on Literacy in Out-of-School Youth and Adults, co-organized by UNESCO and ILI, and held at UNESCO, Paris on June 22-24, 1998. The views expressed in this paper represent those of the author, rather than those of either sponsoring organization. The author would like to express his appreciation to the following individuals for their useful comments on an earlier draft of this paper: Dieter Berstecher, Vina Chinapah, Mohamed Maamouri, Adama Ouane, and Abdelwahid Yousif.

Abstract

The World Conference on Education for All (WCEFA) in 1990 at Jomtien, Thailand, included a number of educational targets related to out-of-school youth and adults, including (a) to reduce the number of adult illiterates to half of the 1990 level by the year 2000 and (b) to improve learning achievement to an agreed percentage of an appropriate age cohort. WCEFA also stressed the need to monitor and evaluate the performance of individual learners as well as the delivery mechanisms and outcomes of literacy and other nonformal education programs. At the Mid-Decade Review meeting on EFA in Amman (in June 1996; see UNESCO, 1996b), the international community further called for efforts at both international and national levels to adopt new techniques and strategies to collect and analyze meaningful data to monitor progress towards the Jomtien goals.

Although many countries have been actively striving to meet the basic learning needs of all, current national and international capacities remain inadequate for assessing and monitoring the acquisition of literacy and numeracy for out-of-school youth and adults. Furthermore, increasing attention to the functional aspects of literacy during the 1990s has reinforced the necessity to assess individual performance across a continuum of literacy abilities not tied directly to school curricula, from the basic 3Rs to the increasingly important area now termed 'basic life skills' or 'survival skills.' There remains a critical need for a practical and low-cost methodology that can bridge the gap between methodologies for assessment at the national survey level. At the program level the methodology must be able to be used effectively by developing countries with limited funds but major literacy problems.

The present paper provides a critical review of the issues raised above, with detailed discussion of (a) cultural dimensions to literacy; (b) international statistics on literacy comparisons; (c) learning, literacy, and the achievement of basic and life skills; (d) national level versus local program needs; (e) methodologies for literacy assessment; and (f) challenges to the validity of literacy assessment tools.

TABLE OF CONTENTS

Introduction 1

Cultural dimensions to literacy comparisons 1

International statistics on literacy 3

Learning, literacy, and the achievement of basic and life skills 6

National level versus local program needs 9

Methodologies for literacy assessment 9

Challenges to validity of literacy assessment tools 11

Conclusions 13

References 16

INTRODUCTION

The Framework for Action adopted by the World Conference on Education for All (WCEFA) in 1990 at Jomtien, Thailand, included a number of educational targets related to out-of-school youth and adults, including (a) to reduce the number of adult illiterates to half of the 1990 level by the year 2000 and (b) to improve learning achievement to an agreed percentage of an appropriate age cohort (e.g., 80% of 14 year olds). WCEFA also stressed in this regard the need to monitor and evaluate the performance of individual learners as well as the delivery mechanisms and outcomes of literacy and other nonformal education programs. At the Mid-Decade Review meeting on EFA in Amman (in June 1996) the international community further called for efforts at both international and national levels to adopt new techniques and strategies to collect and analyze meaningful data to monitor progress towards the Jomtien goals.

Although many countries have been actively striving to meet the basic learning needs of all, current national and international capacities remain inadequate for assessing and monitoring the acquisition of literacy and numeracy for out-of-school youth and adults. Furthermore, increasing attention to the functional aspects of literacy during the 1990s has reinforced the necessity to assess individual performance across a continuum of literacy abilities not tied directly to school curricula, from the basic 3Rs to the increasingly important area now termed 'basic life skills' or 'survival skills.' There remains a critical need for a practical and low-cost methodology that can bridge the gap between methodologies for assessment at the national survey level. At the program level the methodology must be able to be used effectively by developing countries with limited funds but major literacy problems.

Overall, a reliable and practical information base on literacy and basic life skills is required, both within and across diverse national and cultural contexts. The present paper (as related to the Expert Seminar on Literacy Assessment of June 22-24, 1998) provides a review of issues pertinent to the matters raised above, with some proposed suggestions for future steps to action in this domain.

CULTURAL DIMENSIONS TO LITERACY COMPARISONS

Many countries have attempted to better understand national, social, economic, and educational policies in light of other nations' successes and failures. In principle, there is nothing wrong with comparative analysis; indeed comparisons are one of the few ways by which nations can gauge their progress. However, comparative studies can also be, for a variety of reasons, misleading. One way to assure that what are termed cross-national comparisons are credible is to relate them to the long and robust cross-cultural and comparative research tradition. Largely begun by anthropologists in the colonial era, such cross-cultural studies examined and compared human behavior ranging from child rearing practices and initiation rites to the training of craftsmen and sedentary agriculturists. Since the 1950s, work on international and cross-national educational comparisons was undertaken in earnest, as exemplified by a broad array of research under rubrics such as education and development (Fägerlind & Saha, 1983), cross-cultural psychology (Berry, Poortinga, Segall, & Dasen, 1992), and cultural psychology (Stigler, Schweder, & Herdt, 1990).

Cross-national comparisons of literacy acquisition and reading skills are of rather recent vintage, beginning with Downing's (1973) seminal *Comparative Reading* to Scribner and Cole's (1981) well-known research on literacy in Liberia, to the more recent IEA study of reading in 32 school systems undertaken under the auspices of the International Association for the Evaluation of Educational Achievement (IEA) (Elley, 1992). Downing's work focused on the cultural specificities of reading in different languages and scripts, Scribner and Cole's on the cognitive consequences of literacy in Africa, while Elley's study evaluated commonalties of performance across a large number of countries from the industrialized and developing worlds.

We know that literacy is not only difficult to define in individuals and delimit within societies, but it is also charged with emotional and political meaning. It was not long ago that newspapers and scholars assigned to whole

INTERNATIONAL LITERACY INSTITUTE
PENN Graduate School of Education

societies a single referent such as "illiterate and uncivilized"; and, "illiterate" is still a term which carries a negative connotation around the world.

Defining literacy as an individual cognitive ability was once thought to be simple: it entailed the testing of reading and writing skills. This could involve, as in American voting laws of the early 20th century, the ability to sign one's name, or even mark an "X." Or, as is done in some present-day societies for statistical and methodological expediency, literacy may be simply inferred from school attendance: those with four (or eight or twelve) years of formal public schooling are assumed to be literate. Or, in yet other societies, literacy rates are calculated from the numbers of persons who answer "yes" to the simple question "Can you read and write?" It is now known that such approaches to presumed literacy may be quite misleading, for a host of reasons. Furthermore, when considered as a cultural—as opposed to a cognitive—phenomenon, literacy is even less well defined, since its meanings, functions, and methods of transmission may vary greatly from one cultural group to the next.

The distinction between "emic" and "etic" concepts (Berry & Dasen, 1974; Pike, 1966) is a central feature in cross-cultural and comparative studies. *Emic* concepts are those that can be understood only within a single cultural system or society, and are measured only according to criteria relevant to and understood within that system. *Etic* concepts are those that are deduced or derived from a position outside of any particular system, and have as a primary goal the analysis of more than one social system or society.

For example, an etic perspective on literacy assumes that skills such as decoding, word-picture-matching, and reading a bus schedule ought to have substantially the same meaning to different individuals and across different cultural groups. An emic perspective on literacy would encompass skills and meanings associated with literacy within cultural groups, such as "script recognition" skills in different orthographies, as well as the knowledge about the values, meanings, uses, and even attitudes of/about print in everyday social contexts. Simply put, emic literacy skills are those which can only be adequately understood within a given society; etic skills are those that have developed out of the heuristic convenience of those who desire a common frame of reference or system of measurement.

The three approaches mentioned earlier—Downing, Scribner/Cole, and Elley—provide useful examples for examining the problem of cultural comparison. Downing (1973) asked specialists in a variety of languages to analyze the nature of reading development within specific languages and orthographies, such as Hebrew, Chinese, French, Japanese, and English. While most contributors to the Downing volume utilized standard frameworks for reading analysis, such as word decoding and sentence comprehension, they also focused on what was *specific* to the nature of learning to read in the language and social context of the given society.

In contrast, Scribner and Cole (1981) addressed the classic issue of whether there are cognitive consequences to becoming literate, through a combined ethnographic and quasi-experimental study among the Vai people of Liberia. Their methodology focused primarily on the cognitive and literacy skills possessed by non-literate tribal people, by providing within-culture comparisons between schooled and non-schooled adults. Among other things, the researchers were interested in understanding how social practices, such as Quranic study and letter writing among commercial traders in Liberia, would affect both cognition and literacy. The results, integrating emic and etic approaches to data gathering and analysis, provide one of the most in-depth portraits of literacy and culture that is available to date.

More recently, the IEA study of reading literacy (Elley, 1992) employed an etic approach to the study of reading in 32 countries, with the general aim of comparing reading achievement levels, as well as looking for relationships between social variables and cognitive outcomes. Considerable efforts were made to assure the adequate translation of test items, as well as the similarity of the 10- and 14-year-old populations surveyed, but, Elley (1992) states, "(i)n the end,... each reader must decide for him or herself whether a comparison is adequate for a particular purpose," since "comparisons are useful for many purposes and are certainly better than alternative comparisons made without carefully controlled empirical data." While it is difficult to argue against the idea of more (empirical) data, it is less than

certain that "more" always means "better." Indeed, in the information-driven societies of today, there are times when the sheer volume of data can overwhelm rational analyses that are based on direct observations of local reality.

In sum, etic (external, quantifiable, comparison-oriented) measures are clearly important in understanding how people acquire literacy, how educators and policymakers view literacy, and how economic and societal systems interact in an increasingly interdependent world. Throughout the 20th century, literacy has been thought to be associated with formal schooling, student achievement, and economic success. Thus, it is not surprising that quantifiable tools are needed to compare children with other children, adults with other adults, and social and cultural groups with other such groups, depending on policy and programmatic needs. Nonetheless, the psychometric measurement of literacy (especially the form taken in large-scale comparative surveys) is not the only way to view the acquisition or possession of literacy skills, though it may well be the most convenient and cost-effective way for policymakers to measure skills across groups of people.

INTERNATIONAL STATISTICS ON LITERACY

In order to provide worldwide statistical comparisons, UNESCO has relied almost entirely on data provided by its member countries (UNESCO, 1983). These countries, in turn, typically rely on national censuses, which most often determine literacy ability by self-assessment questionnaires and/or by means of a proxy variable utilizing the number of years of primary schooling. Because of doubts about the reliability of such data—especially in the light of rapid social, economic, and demographic changes—considerable concern has been expressed about the credibility of literacy statistics. In 1986, for example, UNESCO and the UN Statistics Office (UNSO) held a joint seminar in Paris to discuss the use of household surveys to improve adult literacy statistics; a technical report which was the basis of this seminar was later published (UNSO, 1989).

While many specialists now agree that exclusive reliance on indirect measures of literacy may be flawed, there is renewed discussion of the utility of proxy measures (Murray, 1997). More importantly, there have been changes in the way that some countries have begun to view literacy measurement, especially in North American national surveys, and in the OECD countries (see discussion of the International Adult Literacy Survey, or IALS, study, 1995, below).

The direct measurement of basic skills in youth and adults, however, can be technically challenging and costly. For these reasons and others, most countries (especially in the developing world) have not as yet chosen to invest in local, national, or regional direct measurement survey strategies. Furthermore, it must be recognized that any change in the methodology used for calculating literacy rates in a population might result in uncomfortable political consequences; for example, not long ago in the United States the 'official' U.S. literacy rate of over 95% was reduced to about 75% of the adult population, as a consequence of redefinition and survey assessment (via the U.S. National Adult Literacy Survey [NALS], Kirsch et al., 1993). Similar adjustments (in a downward direction) were a consequence of the IALS study in 1995, and eventually led to the withdrawal of France from the final report.

Overall, the gathering of international statistics on literacy poses a variety of problems which have been matters of debate among specialists. With the greater need for direct measurement, and increased technical capacity for such measurement in developing countries, some clarification of issues seems to be taking place, as described below.

Language policy and multilingualism

Most countries have formulated an explicit language policy that typically states which language or languages have official status. Often the decision on national or official language(s) is based on such factors as major or dominant linguistic groups, colonial or post-colonial history, and the importance of a given language to the interests of economic development. Official languages are also those most commonly used in primary school, although there may be differences between languages used in

INTERNATIONAL LITERACY INSTITUTE
PENN Graduate School of Education
UNESCO

BEST COPY AVAILABLE

early schooling and those used later on. Further, there may be important differences between language policy in primary schooling and that of nonformal education (NFE) and adult education. For example, in Senegal, French is used exclusively in primary school, but local Senegalese languages are used in adult literacy programs nationwide. The use of mother tongue instruction in primary and adult education remains a topic of debate (Engle, 1975; Wagner, 1992).

Which languages 'count' in literacy statistics is also controversial. When UNESCO gathers literacy statistics, it typically depends on government census bureaus to report the number of literates and illiterates, as well as age and gender differences. Apparently, the language of literacy is not always specified by the countries concerned. That is, it would be difficult to know, taking the case of Senegal again, how many 'literate' individuals are literate in French, Wolof, Arabic, or other languages, and how this is broken down by age and gender.

While there is usually general agreement that all official language(s) ought to be assessed in a national literacy survey (e.g., English in the United States; English and French in Canada, and German, French, Italian, and Romanch in Switzerland), there may be disagreement over the assessment of literacy in non-official or semi-official languages, where these have a recognized and functional orthography (e.g., Athabaskan in Canada, or Hungarian in Romania). In many countries, there exist a multitude of local languages that have varying relationships and status with respect to the official language(s). How these languages and literacies may be included in a national literacy survey can be a matter of serious debate. For example, in certain predominantly Muslim countries in sub-Saharan Africa (e.g., Senegal or Ghana), the official language of literacy might be French or English, while Arabic—which is taught in Islamic schools and used by a sizable population for certain everyday and religious tasks—is usually excluded from official literacy censuses. Similarly, literacy in Chinese, Spanish, Cherokee, and other written languages have generally been ignored in literacy assessments in the United States.

Emic and other perspectives would suggest that all languages and scripts should be included in national literacy assessments in order to fully understand the range of human resources (and to maximize human rights). Practically speaking, this poses various problems in fiscal and technical resources. However, assuming sufficient support for undertaking a household literacy survey, effort should be made to investigate and assess literacy in all "significant" language populations, where significant might be defined by population size or percentage (see Wagner, 1990). This is essential especially because problems of illiteracy and low literacy tend to be over-represented in linguistic minority groups. Further, for comparative research, studies which include the diversity of the languages and literacies within a country provide a more accurate measure of social variation than those with a more limited focus.[1]

Classification issues

In the 1960s and 1970s, when many developing countries entered the United Nations, the majority of the youth and adult populations of these countries typically had never gone to school nor learned to read and write. It was relatively easy in those contexts to define such individuals as "illiterate." The situation now is more complex, as some contact with primary schooling, nonformal education programs, and the mass media is made by the majority of families across the globe, and now even in the poorest of communities. Thus, even though parents in many countries may be illiterate or nearly so, it is not unusual for one or more of their children to be able to read and write to some degree. For this reason alone (and there are other reasons, as noted below), it would seem that simple dichotomous classifications—still in use by many international organizations and national governments—ought to be replaced by measures that better represent the range or continuum of literacy abilities which are common to contemporary societies.

The advent of the WCEFA in Jomtien, with its call for an end-of-decade assessment in the year 2000, and the growing use of surveys of learning achievement in the 1990s, there is increasing awareness of the utility (and challenges) of household survey methods

[1] Nonetheless, it should be noted that it is not uncommon for governments to make arbitrary language and literacy "exclusions" from nationally representative samples for reasons of efficacy, politics, ethnic strife, and so forth.

INTERNATIONAL LITERACY INSTITUTE
PENN Graduate School of Education
UNESCO

for measuring literacy levels. In such surveys, levels of literacy achievement are ascertained based on observed literacy skills. These data are then used to perform statistical analyses in relation to other variables (such as gender, age, schooling, language background, and so forth), as discussed below.

International comparability of data

The comparability of data—across time and countries—is a major concern for policymakers and planning agencies. If definitions and classifications vary, then it can be difficult if not impossible to compare data collected through different surveys. Comparability and stability are, for example, the hallmarks of the UNESCO World Education Report and UNESCO's new World Education Indicators project. Nonetheless, if comparability is the primary goal, then less attention is paid usually to the local and cultural validity of the definitions and classifications of literacy; hence data may become less meaningful or potentially misleading.

International and national needs, definitions, and research strategies may or may not come into conflict over the issue of comparability, depending on the particular problem addressed. For example, as mentioned above, UNESCO solicits literacy data worldwide, where literacy is measured in terms of the number of "literates" and "illiterates." For most countries, this dichotomous type of classification presents few practical (technical) problems and is relatively inexpensive to gather (as part of nationwide censuses), while providing international agencies with a cross-national and time-series framework for analyzing literacy by geographic or economic world regions. However, the fact that this dichotomous literacy variable may be regarded as simplistic and flawed, places serious limits on its usefulness in comparative statistics.

For example, educational planners may want to know about the effects of the completion of primary or secondary schooling—or of a particular literacy campaign—on levels of literacy attainment. In these cases, a simple dichotomy is a blunt statistical instrument; literacy scores or levels are clearly required for learning achievement to be adequately measured. Furthermore, precise data are needed as to which languages and which literacies are used in each region and by ethnic group, in addition to age and gender variation.

Compared to census-based literacy estimates, household-level surveys offer considerable opportunity to create a detailed picture of literacy skill profiles and their demographic correlates in national populations. As detailed in the IALS survey, such surveys can achieve considerable cross-national comparability, but not without various problems related to item comparability (Levine, 1998), and population sampling (Kalton et al., 1998).

The impact on educational policy of such comparative studies is due at least in part to the realization among national policymakers that their national or regional levels of learning achievement may be far lower than would have been predicted, say, by participation in school. The 1994 U.S. National Adult Literacy Survey found, for example, that 50% of minority group adults in America with a 10th-grade education could read at only a 5th-grade level. Issues of equity in the United States continue to dominate educational policy, as minority youth are, on average, about 4 years behind achievement levels of same-age White children, even up through college years (a Black college graduate in the United States reads at about the level of a White high school graduate). In a recent World Bank national household survey in Bangladesh, Greaney (Greaney et al., 1998) found that five years of primary schooling resulted in only a first grade equivalent of learning achievement, and that three years of schooling had approximately zero value in terms of learning achievement. This study may have an important impact on the kinds of investments that Bangladesh makes in the areas of basic and nonformal education in the future.

Measuring the consequences of literacy

Will a change from a lower to a higher level of literacy skill make a concrete difference in an individual's life? Looking at the "average literacy rate" and comparing this statistic with health indicators (such as infant mortality rates or fertility rates), or estimating "employability" from such a rate, cannot adequately illuminate the diversity of individual human conditions. Nonetheless, the author's own work in rural and low-literate African countries

has demonstrated that those with higher literacy tend to be better off economically (Wagner, 1993); and the IALS found similar evidence across OECD countries.

Since the 1950s, perhaps the most compelling argument for human resources development is that literacy and schooling will lead to economic growth in countries that make a sufficient and appropriate investment in them. This is the approach sometimes referred to as "investment in human capital" (Psacharopoulos & Woodhall, 1985). Bowman and Anderson (1973), for example, went so far as to claim that an 80% national adult literacy rate would be necessary for rapid economic development, while a 40% literacy rate would be required for a minimum take-off of economic development. This type of claim makes use of aggregated data across many countries of the world, based on a significant correlation between gross national product (GNP) and literacy rates (using UNESCO data). The notion of causality imposed on such correlations is very hazardous, but it is still quite common in the discourse of economic and literacy planners. In one of the few studies of economic impact, it was found that, in the agricultural sector, an additional year of primary schooling directly affected wages and farm output (Jamison & Moock, 1984) even though literacy achievement itself was not measured.

The issue is not so much whether, but rather *how* to promote literacy in a way that is consistent with overall policy objectives. How can one most efficiently achieve increased literacy levels (and important social and economic impacts) within the available economic and social resources? Here, the emphasis is on efficiency and appropriateness, domains in which specialists—with appropriate measures of learning/literacy achievement—can play a crucial role. What is efficient and appropriate, however, brings forth a whole series of qualifiers, such as for whom, in which language, for what purpose, using what methods, and so on.

There is also the issue of program impact, at the local level. Various agencies have made serious investments in NFE and literacy programs in the context of their development strategies for women, youth, and disadvantaged adults. The measurement of impact of these programs not only on learning achievement, but also on social and economic outcomes, has become a matter of some urgency. The development of low-cost measures of assessment in this regard might involve not only learning achievement but also other indicators (such as attendance, attitude, and so forth) that pose special challenges, from technical and statistical perspectives (see below for further discussion on program level assessments). Finally, there is also the important question of literacy (or learning) retention—how much learning is retained by learners for use after program participation ceases (Wagner, 1998).

LEARNING, LITERACY, AND THE ACHIEVEMENT OF BASIC AND LIFE SKILLS

The WCEFA took a dramatic step by supporting the notion that basic education consists of more than conventional (or formal) schooling—that education for all should include a set of knowledge, skills, and values that serve as the basic learning competencies (see below). The EFA declaration also made clear that learning achievement (rather than participation rates, diplomas, etc. that predominated in the past) should be the primary focus of action.

In addition, the EFA approach put the emphasis not only on measuring learning achievement (rather than simply monitoring indicators such as expenditures, numbers of graduates and teachers, and so forth), but also on widening the lens of view to beyond the 3Rs. Such an approach raises a number of definition issues and questions, as discussed below.

Definition issues and questions

Literacy.

Many definitions exist for literacy. Two well-known and oft-cited definitions are as follows:

> A person is literate who can with understanding both read and write a short simple statement on his everyday life.... A person is functionally literate who can engage in all those activities in which literacy is required for effective functioning of his group and community.... (UNESCO, 1978)

BEST COPY AVAILABLE

> Using printed and written information to function in society to achieve one's goals and to develop one's knowledge and potential. (OECD/Statistics Canada, 1995)

While a comprehensive review of these and other definitions cannot be provided here (see Venezky et al., 1990, and Wagner, 1992), there is nonetheless substantial consensus that literacy skills refer to the cognitive abilities of reading, writing, and calculating. How to measure these abilities (and the contents of those measurement tools) is still widely debated. Whatever tools will be used to measure literacy will likely have to include measures of these three skills.[2]

Learning and life skills.

Article 1 of the WCEFA Declaration states that basic learning needs or competencies

> comprise both essential learning tools (such as literacy, oral expression, numeracy and problem solving) and the basic learning content (such as knowledge, skills, values, and attitudes) required by human beings.... (UNESCO, 1990, p. 43)

Thus, following Jomtien, learning may be seen not only in terms of the 3 Rs, but also in terms of the kinds of skills, competencies, knowledge, and even attitudes relevant to human development. That is, one should focus both on the skills and contents of the instructional 'package' — on what are the purposes of learning (and, by extension, of the education provided).

This definition of basic learning competencies is extremely broad, and hence poses serious problems for those who wish to use it as a focus of either planning or assessment. Recent attempts to consider learning of 'life' or 'survival' skills have received increased consideration in recent years, but problems of culturally appropriate content pose serious concerns. One person's survival skills may be radically different from another's, and these vary not only across cultures, but perhaps just as significantly within countries. One illustrative example of this problem is the current development of an international life skills survey by the OECD, as a follow up to the IALS study. 'Technological literacy' skill is part of this prospective study; yet, one can easily imagine that poor sectors of developing countries might find such a measure to be quite out of place at this moment in time. Nonetheless, there have been attempts within schooling achievement studies to develop measures of life skills (see MLA, Chinapah, 1997, discussed below).

Learning Achievement.

Learning achievement is understood here mainly in terms of what can be measured through direct assessment (rather than through proxy variables). But it may also be seen in terms of who is the end-user, and what are the timeline constraints of data gathering; that is, what purpose(s) does this measurement of learning achievement address? Who will use this information, and for what purposes? Some end-users (such as service agencies, including many NGOs) are uncomfortable with evaluation and monitoring in general, and with learning achievement more specifically. Their anxiety is often represented by the agency's focus on service delivered in terms of educational participation, rather than impact on learning. In other words, learning achievement and skill assessment are not neutral topics, and may be subject to widely differing approaches, which development or national agencies will need to choose amongst for their planning and programming efforts

Clientele(s).

Who are the real clients of education and development (learning and literacy) programs? Beyond the 'learners' themselves (children, youth, and adults), clienteles might also include the following: children (of learners), families, communities, political constituencies, program staff, government agencies, the public, and others. Are these clients all given the same valence, or are some more determinant of policy than others, and why? For example, if

[2] Calculating—within the broader domain of numeracy—while often ignored in the past in developing country NFE programs is important to include in future work on literacy assessment. But including numeracy can be misleading in some cases: where applied in NFE programs in developing countries, numeracy has unfortunately been biased often toward school-based learning, an approach which is likely to be less than appropriate for those in NFE programs and who use math skills in everyday and work activities.

INTERNATIONAL LITERACY INSTITUTE
PENN Graduate School of Education
IIEP

BEST COPY AVAILABLE

educational access for the 'unreached' (or least-reachable) is a policy priority, then the broad category of 'out-of-school youth' as client may require more specification—which youth, what gender and ethnicity, what level of poverty and distance, and so forth.

Examples of the monitoring of learning achievement

In the domain of learning achievement, there are a range of studies—some national, others international—that have focused on the achievement of reading, math, science, and so forth, both in school and out. For example, the IEA (International Association for the Evaluation of Educational Achievement) undertook a number of important international comparative studies of learning achievement, such as the 1992 Reading/Literacy study (Elley, 1992) of 9- and 14- year-olds in 32 countries (including a number of developing countries such as Zimbabwe, Nigeria, and Botswana), and the 1996 Third International Math/Science Study (TIMSS) of 4th- and 8th-grade students in 26 countries (including developing countries such as Thailand and Iran) (U.S. Department of Education, 1996). The 1995 IALS study, in 6 OECD countries, measured out-of-school reading, writing, and math skills in adults aged 16-65 years. Each of these studies contained somewhat parallel methodologies for the measurement of learning achievement, such as the use of item sampling pools, translation-retranslation methods, and psychometric validation for item rejection.

In international/comparative studies, factors such as time-on-task on a particular school subject (such as math in the TIMSS), or the impact of TV or book availability (IEA reading/literacy study) on achievement can be studied and disaggregated through multifactoral statistical techniques. In this manner, it was found, for example, that the relatively high amount of time spent by U.S. students in classroom work in math was not a likely cause of high math achievement, as several other high-scoring countries spent relatively lesser amounts of time studying math in class. Similarly, it has been shown that smaller class size is not strongly determinant of high achievement, as some countries (e.g., Japan) have large class size and high achievement scores. For a number of reasons, such conclusions have been the subject of serious criticism, particular due to the inability of large statistical analyses to deal with local cultural explanations. For example, large class size in Japanese schools may be less of an impediment to learning because of the strong socialization processes that reinforce group solidarity and close-knit behavioral norms.

While there is legitimate interest (and sometimes significant amounts of funding) for such international comparative studies, there has been less interest (and less funding) for efforts to measure basic learning achievement at the program level. This observation appears particularly obvious in LDCs, where NGO-based programs rarely have the capacity to engage in empirically sound self-evaluation, and where international agencies infrequently have the resources (human or fiscal) to invest in evaluations that include learning achievement (there are some well-known exceptions to this generalization, such as BRAC in Bangladesh or TOSTAN in Senegal). With the risk of possible oversimplification, it seems fair to say that many (if not most) of the innovative educational programs sponsored by development agencies in collaboration with NGO and/or GO agencies, seldom have the benefit of formative or summative evaluations which include learning achievement; nor, with few exceptions, have such programs invested in local capacity building.

Three projects that have received support from UNICEF in the domain of learning achievement are worth mentioning in this context: the ABC (Assessment of Basic Competencies) project in Bangladesh (Unicef, 1992), the MLL (Minimum Levels of Learning) project in India (NCERT, 1995), and the MLA (Monitoring of Learning Achievement) UNESCO-Unicef project (Chinapah, 1997) in more than a dozen LDCs. Each of these projects contains interesting dimensions. For example, the ABC project represents an attempt to design an evaluation methodology that is exceptionally low in per person assessment cost. The MLL project attempts to take a fresh look at what kinds of skills can and should be measured both in and out of school. The MLA project tries to link monitoring methods to training and capacity building in diverse countries in order to develop new ways of promoting a "monitoring culture."

NATIONAL LEVEL VERSUS LOCAL PROGRAM NEEDS

The field of basic education has, over time, been subject to a number of artificial separations or 'walls' between sectors such as formal and nonformal, and for children, youth, and adults. Bureaucracies and administrative needs tend to reinforce such separations, but the contents of the learning (and assessment) of basic skills need not and should not maintain such walls. Indeed, it is important to utilize available resources across these sectors of activity. In a similar way, there has been a tendency, in the gathering of literacy data, to think either in terms of national level needs or in terms of local program (or project-based) needs. Below are described the goals and tools typically employed in these two domains.

Goals

The issues and problems of gathering international statistics on literacy fall generally into the domain of 'national level' data gathering—whether through traditional censuses and proxy measures, or through household sample designs. The general goal of national level data collection includes at least the following practical purposes: comparison of literacy rates over time, determination of statistical relationships between literacy and other variables of interest (e.g., gender, years of schooling, etc.), and cross-national comparability.

However, there are other major assessment needs in literacy work besides those listed above. Most of these would fall under the rubric of program (or project) impact, planning, and development at the local or regional level. The specific purposes of assessment and evaluation at the program level include the following: comparison of program effectiveness over time; comparison of one type of program with others; effectiveness of teacher training; and so forth. The emphasis in program assessment and evaluation is on 'how programs work,' rather than on national policy planning at the national or global levels. At the program level, there are real and frequent needs to know what kinds of learning achievement impacts are being realized.

Tools

In the broad domain of nonformal basic education, adult education, and literacy education, there has been a tendency to assume that the two broad measurement needs—national and local/program—would use quite different tools or methods for assessment. Indeed, this has most often been the case in the past, at least in part due to the difficulty of finding sufficient technical and fiscal resources at the local level to mount impact studies. There is no particular reason for this situation to remain unaltered. Indeed, there are major gains to be made if the methodologies and purposes of national and local assessments were conflated, particularly in the sense of shared perspectives, tools, and human and technical resources.

In sum, there is a considerable need to take a fresh look at the complementarities between national/international and program level information needs, using surveys of direct measurement of learning.

METHODOLOGIES FOR LITERACY ASSESSMENT

As noted above, there have been many efforts to measure literacy achievement over past decades. Given the diverse stakeholders in the literacy and learning achievement fields, there exists a wide variety of tools for assessment; but there is also some confusion with regard to the terminology employed to describe such methods.

Terminology in literacy assessment

A brief list of key terms are as follows, with a short summary of how the term is typically used:

Assessment.

Assessment refers to one or more methods for judging (via some type of empirical observation or test) the actual performance of literacy or other cognitive skills. For example, the UNSO (1989) and IALS (1995) surveys included literacy assessment, as did the household survey of basic learning achievement in Bangladesh (Greaney et al., 1998). Of course,

with the wider definition of WCEFA mentioned above, assessment might have to take into account broader areas of knowledge, attitudes, and even values, although consensus on assessment tools in these areas remains much more problematic (cf. Little & Wolf, 1996).

Monitoring.

Monitoring refers to one or more methods (sometimes referred to as indicators) for gathering and analyzing data related to (but not necessarily including) the measurement of literacy or other cognitive skills. For example, monitoring can refer to data on rates of schooling or literacy program attendance, diplomas attributed, and so forth, as included in a wide variety of literacy evaluation reports (e.g., UNESCO, 1996a). The UNESCO-UNICEF project on Monitoring Learning Achievement (Chinapah, 1997) employed indicators as well as tests of learning achievement.

Survey.

Surveys typically gather a wide range of information, qualitative and/or quantitative, and may or may not include performance assessments (tests). Literacy surveys, such as the UNSO and IALS surveys, gathered background information as well as test data on individuals.

Evaluation.

Evaluation studies (in this case, 'summative evaluation') refer to attempts to judge the intended impact of any program of activity. In literacy work, evaluations have usually taken the form of qualitative or quantitative accounts of a given program. In practice, this might include the level of learner and teacher participation, the attitudes of learners, the number who 'successfully complete the program,' and so forth (see examples in Bhola, 1990; Easton, 1997). Reliable achievement data has rarely been gathered in LDCs in literacy programs.

Census-based literacy surveys

Since its founding, UNESCO has gathered literacy statistics from its member states that are almost entirely based on the census surveys (mostly on a 10-year [decennial] basis, with occasional updates). Given all of the usual limitations of such surveys in terms of time available and competence of census survey staff, it is rare that literacy performance measures are employed. Thus, most census survey information used to calculate literacy rates employs either proxy measures (mainly the highest level of schooling reached) or simple questions of self-assessment (e.g., "Are you literate?"). Little further specification is obtained in such census-based surveys. Hence, such surveys are often out of date (given the rapid changes in many countries) and very approximative (given the low reliability of the data collected).

Household-based surveys

Household-based surveys offer considerable potential in the field of literacy and basic skills learning achievement. This is the case in large part due to the many options for flexibility in design, such as population sampling, cognitive knowledge sampling, and cost management (cost can be traded off against gains in statistical reliability). For other trade-off issues, see Vrignau and Bonora (UNESCO/ILI, 1998) and Kolstad (1997).

For example, with a focus on developing countries, Wagner and Srivastava (UNSO, 1989) produced a relatively low-cost method for the development of household surveys of literacy, based on a national household survey in Zimbabwe. While the analyses of the Zimbabwe survey were never fully completed, aspects of the methodology were adapted by others for household survey use, such as in Zieghan (1992) and by the World Bank in Ghana (Glewwe, 1991), in Morocco (Lavy et al., 1995), and in Bangladesh (Greaney et al., 1998). Each of these studies utilized a national (as opposed to an international-comparative) framework to produce items that could, with reasonable reliability, assign individuals to levels of literacy and numeracy achievement. Multiple languages were used in the UNSO study, with a rudimentary attempt at psychometric comparability (i.e., more could have done in this regard if more resources had been available).

The IALS survey (OECD/Statistics Canada, 1995, 1997; and its predecessors, such as the 1993 U.S. National Adult Literacy Survey (NALS), invested significant resources

in improving the technical and psychometric properties of literacy assessment instruments, using a variety of techniques, including methods for expanding the range of items used in a survey format (including especially Item Response Theory or IRT). The IALS, and its predecessors, utilized a five-level categorization method for literacy, along three different scales: prose literacy, document literacy, and quantitative literacy (or numeracy) (for an overview of this methodology, see Tuijnman, Kirsch, & Wagner, 1997). These survey scales are not without critics (e.g., Reder, in press, on the NALS and on the colinearity of the three scales); and there are criticisms, for example, on the degree of international comparability (e.g. Kalton et al., 1998, on population sampling difference across IALS countries) or on item comparability (Levine, 1998).

Such international surveys have also been criticized for being too expensive and too complicated to be 'owned' (that is accepted for endogenous and locally sustainable use) by national and subnational agencies. Nonetheless, these comparative surveys have often received considerable media and policy attention, and at times have led to significant national educational policy shifts. The IALS, for example, has led the United States, Canada, and a number of other OECD countries to invest additional resources in adult literacy and adult education programs. National or subnational household surveys can also have a similar policy impact (as in the U.S. NALS), but this result necessitates a serious and credible study, followed by concrete efforts to publicize results.

Program evaluation studies

Both formative and summative evaluations of literacy programs in developing countries are numerous (Bhola, 1990; Carron et al., 1989). In developing countries especially, such evaluations have only rarely included psychometrically appropriate tests for measuring learning achievement. Typically what is measured are the inputs in human and infrastructural resources, the pedagogical methods employed, and the outcomes in terms of attendance and 'successful program completion.' In industrialized countries as well, there is insufficient attention to the impact of programs on learning achievement (Wagner & Venezky, 1998). This is a very serious gap in the current knowledge base.

Cost issues

Costs of learning achievement and assessment studies are quite variable. Estimates of the total cost of the IALS survey(s) run as high as tens of millions of U.S. dollars, while the costs of the UNSO national literacy survey in Zimbabwe may be estimated at about US$100 thousand in current dollars. Costs clearly depend on the number of countries included, the degree of external expertise required, the complexity of study design, and degree of collection and analyses undertaken. Clear tradeoffs are available in costing processes, from limiting sample sizes to the length of tests created to the degree of trained personnel required. Nonetheless, there is little exactitude in current cost estimation due to the paucity of available studies with cost figures.[3] (This is also true with respect to costs of literacy programs themselves, as described in Wagner, 1995.)

CHALLENGES TO VALIDITY OF LITERACY ASSESSMENT TOOLS

All measures of human behavior are subject to what have been termed 'challenges' to their validity. This is nowhere more the case than in literacy education. In part, this is historical, in that literacy work has been often tied to politics, and thus not subject to the ordinary checks and balances of social science research and evaluation studies. Furthermore, since literacy has been seen over time as a skill that can 'easily be taught' (for example, in 'each one, teach one' campaigns), it is only relatively recently (the last two decades) that the efficacy and efficiency of literacy campaigns and programs have come into question (and more recently, the same being true of formal schooling as well).

[3] According to Greaney & Kellaghan (1996) the typical cost components of (national) assessments are rougly as follows (as percent of cost): instrument development 15; sampling and selection 10; data collection 30; data processing 10; data analysis 15; reporting and dissemination 15; and governance 5.

INTERNATIONAL LITERACY INSTITUTE
PENN Graduate School of Education

BEST COPY AVAILABLE

Cross-cultural comparisons

If literacy is so deeply tied to the integral definition of separate cultures, and if a broad definition such as that proposed by the IALS is used, then how can common understandings of literacy levels or competencies be reached across cultures? That is, literacy in poor communities may be viewed (and, in practice, used) in quite different ways than is the case in wealthy communities (see Levine's critique of the IALS, 1998). Indeed, the definition of 'poor' may well have a radically different operational definition in different societies, and often does. Further, when one attempts to compare individuals or groups on whether they are self-defined as "literate," it can emerge that different definitions preclude valid comparison. This is often the case with census-derived information where large differences in schooling bias the decision matrix for responding to the question, "Can you read and write?" (cf. Lavy et al., 1995).

Targets and 'minimum' levels of achievement

The WCEFA Framework for Action states that:

> Basic learning needs...vary...from country to country depending on the actual status of basic education... (C)ountries may wish to set their own targets, such as a... (r)eduction of the adult illiteracy rate...to say, one-half its 1990 level by the year 2000... (Furthermore) setting levels of performance and the proportions of participants who are expected to reach these levels...must be an autonomous task of individual countries. (UNESCO, 1990, Articles 3, 8, and 9, pps. 52-53)

Since Jomtien, policymakers in many countries have given high priority to the improvement of literacy and basic education in low-literate societies or in low-literate 'pockets' of industrialized societies. Yet only a few countries have focused on the minimum levels of basic learning competencies implied in the WCEFA Framework (see the MLL study in India in NCERT, 1995). Further, while a number of countries have reinforced their efforts to try to reach the 'one-half' reduction in illiteracy rates called for by the Framework, there is little evidence that the countries most in need have come anywhere near this target, or would have the requisite data to support their conclusions. Meeting such targets in terms of basic learning needs is even more difficult at the national level as there is often less consensus as yet on what those targets should be.

There are alternative (and more controversial) policies for action (beyond simply more investment). For example, resources might be invested in a more selective fashion (directing more funds to preschool and primary schools, or to specific groups of adults), so that some individuals—those with a greater chance of success—would have access to the possibility to become literate or more literate. Indeed, recent evidence on volunteer literacy efforts in the United States suggests that the least literate portion of the population is remarkably resistant to literacy training, often exhibiting much higher rates of program attrition and lower learning achievement levels (Venezky, 1992). Research in developing countries in this domain could be very illuminating, and might result in quite different conclusions.

Technical/statistical challenges

There are numerous technical challenges to any method of testing. These range from disputes about theoretical and operational definitions to the type of statistical tests employed to analyze data. Given space limitations, it is useful to simply note some areas that have been particularly debated in studies involving literacy assessment. These include the following: (a) scales of literacy achievement (from dichotomous, to 4 levels, to many levels); (b) determination of when a 'level' is achieved (e.g., in IALS, it is adequate to say that a level is achieved if and only if 80% of the items in a level are completed successfully; see Levine, 1998; Sticht, personal communication, 1996); (c) what is included in the operational definition of literacy; and (d) effectiveness of the use of proxy measures (Lavy et al., 1995; Murray et al., 1998; and others).

CONCLUSIONS

In the popular media and in academic circles, it is not uncommon to hear that low levels of literacy and basic skills are among the chief problems facing contemporary societies today. And, there are numerous arguments that would support concern for such a point of view, ranging from the economic pressure on the workplace, to increases in women's health and in farmer productivity in developing countries, and the major problems endemic in urban secondary schools in Europe. Literacy and basic skills development are part of policy discussions in all of these areas, yet needed data that would help to empirically address key policy issues have often been lacking. Improved methods of literacy assessment in out-of-school youth and adults can help fill this gap and provide new and useful directions for future work in this domain.

Choosing the best investment

It is clear that national and international needs may not be one and the same. Countries and cultures are diverse, each with a multiplicity of groups that vary along ethnic, linguistic, social class, economic, and other dimensions. Each country has its own special history of sociopolitical development, and its own experiences with formal schooling and broader educational development. The international policy community has its interests as well, mostly in trying to guide national decision-making from indices of where nations "fall" on some scale of economic productivity or worker efficiency. Thus, the improvement of literacy assessment in a comparative context may affect local, national, and international interests in contrasting ways. National interests and "internal" considerations (involving, for example, population diversity) may be seen as nettlesome problems or simply constraints by planners concerned with international comparison. On the other hand, national considerations about population diversity, linguistic variations, and even orthographic diversity (such as unusual features of a script) may have to be sacrificed in order to achieve a larger basis for international comparison. For these and other reasons, there is ample logic for local programs and national-level policymakers to hesitate in supplanting local concerns and interests for those of the regional or international comparisons.

More specifically, the level of resource investment in empirical data gathering in IEA-like studies is, for many developing countries, far greater than that made heretofore. Thus, there may be opportunities to create a research infrastructure through the carrying out of international comparisons, not so different in kind from that achieved by anthropologists working with diverse groups in developing countries. Perhaps most importantly, comparative studies can, if properly designed, help to achieve a greater understanding of cultural values and attitudes towards learning and literacy. Such an approach would enable the problem of diversity to be turned into an opportunity of studying varieties of learning in context.

What then can be taken from "horse race" comparisons (implicit in IEA- or IALS-like studies), where one country edges out another by a few questions on a test? It would seem that those comparisons are of less immediate interest than what can be learned from cross-cultural and within-culture variations. Unfortunately, sponsors (i.e., funding agencies) of comparative studies usually need the competitive (horse race) aspects to 'sell' the international study to policymakers. Hopefully, a better understanding of the opportunities embedded within cross-national comparisons can lead to a useful change in this funding discourse.

New directions in literacy assessments: Asking the right questions

To undertake program (local) as well as national and international assessments requires an unusual degree of sensitivity to sociocultural variation as well as to the complexities of psychometric and evaluation techniques. The above discussion necessarily raises questions that will need to be addressed in order for serious progress to be made in the domain of assessment of literacy and learning achievement. Certain questions enumerated below have received attention over the years, but remain without broadbased consensus for the moment; others are relatively new in policy discussions.

BEST COPY AVAILABLE

Determining basic learning skills.

What is contained in the set of basic learning skills said to be commonly needed in basic education programs in developing countries? Beyond the 3 Rs, little consensus has as yet been achieved on this question, even though the WCEFA and many of those working in the field seem to have a sense of what 'basics' are needed. While examples of how to engage the process of developing core (basic skills) competencies exist (see earlier discussion of MLA project), they do not provide sufficient clarity as to how a core set of life skills can be broadly applied or agreed upon, especially in out-of-school settings. In other words, what types of skills must be acquired in order to guarantee 'basic learning needs'? This question requires greater conceptual consensus, but also empirical testing beyond the rhetoric still endemic in various policy discussions.

Learning at the program level.

How many educational (and social/economic) development programs actually measure the quality of the program in terms of learning achievement and other related variables? Using what methodologies, and over what time period? For example, in a health education program concerning HIV for young women, what types of knowledge, skills, or behavioral changes would constitute adequate learning achievement? How could these measures be linked to national level a nd household survey data collection? One useful direction would be to have agencies collect an inventory of learning that is expected within programs whose main emphasis is not strictly educational (but includes significant learning components).

'Shareability' of assessment tools.

How can both comparability and context sensitivity be appropriately balanced in literacy and basic skills assessments? Can the tools developed for household surveys be 'shared' with program evaluation and vice versa? What kinds of savings can be made in human and fiscal resources across these two domains? The notion of shareability may be one way to find a compromise between etic approaches to strict (international) comparability on the one hand, and emic approaches to strict national, subnational, and local contextuality on the other.

Timeliness.

Until recently, national literacy data could only be drawn from decennial censuses. More timely data is clearly needed, given the rapidly changing economies and societies in every region of the world. A major question is how can a model of 'continuous' assessment be put into place for both the program and national levels? Are there methods for undertaking reliable household sample surveys that are cheaper, more accurate, and more timely than census data?

Where does learning happen?

Within a learning achievement perspective, the oft-cited distinction between formal and nonformal education may be seen to be of declining importance. Similar knowledge and skills can be instructed and learned whether in or out of schools, whether with children, youth, or adults. If the focus is on learning, then at least one goal would be to better understand which contexts for learning provide optimal access, quality, and relevance. In other words, where do children, youth, adults, and families most effectively learn what they need to learn?

Who does literacy assessment, and the problem of capacity building.

Assessments necessarily require a variety of expertise. The providers of this expertise include policymakers, psychometricians (test makers), survey and interview specialists (enumerators), and data analysts, as well as learners and instructors for preparatory pilot testing. Capacity building—the training of sufficient expert staff—has been a major impediment in the past. How such resources can be put into place in the future will be a major question to address, at local, national, regional, and international levels.

INTERNATIONAL LITERACY INSTITUTE
PENN Graduate School of Education
UNESCO

BEST COPY AVAILABLE

New directions for action: Some next steps

Based on such questions, agencies concerned with literacy assessment (and basic learning needs) will likely have to consider a number of steps in the near and medium term in order to develop a proposed action plan in this domain:

- Initiate a series of several national pilot studies on literacy assessment and basic learning achievement, employing methodologies of use across the program- and national-level needs. In the form of household surveys, these studies would entail program-specific (and learner-specific) observations of what types of learning take place in school, and out of school, as well as in health, agriculture, and the like.

- Develop low-cost and quick methods for the measurement of literacy and basic learning achievement that might be applied to a broad range of education and development activities, and that take advantage of the notion of 'shareability' between both national- and program-level data needs.

- Find ways within development and national agency activities to link learning achievement in education and other sectors (e.g., health, nutrition, agriculture, work, etc.) with the measurement of basic learning competencies.

- Develop a near-, medium-, and long-term plan for capacity building in the area of literacy and learning assessment. Training needs to be reinforced at all levels where reliable data is expected to be collected, analyzed, and reported.

Final thoughts

The task of rethinking literacy assessment will require a better delineation of what literacy and basic (and life) skills mean, how they can be measured, and how knowing more about learning achievement can alter (and improve) policies and practices at all levels of intervention.

By better understanding and avoiding the simplicities of the past, more progress can be made in the future. There still appear to be many educational development projects with learning components in which the impact seems to remain primarily based on conjecture or optimism. This situation needs to change in development work, but how it will change, and what roles different agencies will play, remain less certain.

None of the above activities will be successful without the effective integration of local and national capacity building, which needs to be an integral part of any future effort in this regard. The resulting plan of action should be able to enhance not only national-level assessments but also provide important impact data for programs seeking self-improvement. In this manner, national and international agencies in literacy and basic education will have more effective tools at hand for this central domain of human development. ■

INTERNATIONAL LITERACY INSTITUTE
PENN Graduate School of Education

REFERENCES

Berry, J. W., Poortinga, Y. H., Segall, M. H., & Dasen, P. R. (1992). *Cross-cultural psychology.* New York: Cambridge University Press.

Berry, J. W., & Dasen, P. (1974). Introduction. In J. W. Berry & P. Dasen (Eds.), *Culture and cognition: Readings in cross-cultural psychology.* London: Methuen.

Bhola, H. S. (1990). *Evaluating 'Literacy for Development': Projects, programs and campaigns.* Hamburg: UIE/DSE.

Bowman, M. J., & Anderson, C. A. (1973). Human capital and economic modernization in historical perspective. In F. C. Lane (Ed.), *Fourth International Conference of Economic History.* Paris: Mouton Press.

Carron, G., Mwiria, K., & Righa, G. (1989). *The functioning and effects of the Kenyan literacy program.* IIEP Research Report No. 76. Paris: UNESCO/IIEP.

Chinapah, V. (1997). Handbook on monitoring learning achievement: *Towards capacity building.* Paris: UNESCO.

Downing, J. (1973). *Comparative reading.* New York, NY: Macmillan.

Easton, P. A. (1997). *Sharpening our tools: Improving evaluation in adult and nonformal education.* Hamburg: UIE.

Elley, W. B. (1992). *How in the world do students read?* Hamburg: International Association for the Evaluation of Educational Achievement.

Engle, P. L. (1975). Language medium in early school years for minority language groups. *Review of Educational Research 45,* 283-325.

Fägerlind, I., & Saha, L. J. (1983). *Education and national development: A comparative perspective.* New York: Pergamon Press.

Glewwe, P. (1991). *Schooling, skills, and the returns to government investment in education: An exploration using data from Ghana.* LSMS Working Paper No. 76. Washington, DC: The World Bank.

Greaney, V., & Kellaghan, T. (1996). *Monitoring the learning outcomes of education systems.* Washington, DC: World Bank.

Greaney, V., Khandker, S. R., & Alam, M. (1998). *Bangladesh: Assessing basic learning skills.* Washington, DC: World Bank.

Jamison, D. T., & Moock, P. R. (1984). Farmer education and farm efficiency in Nepal: The role of schooling, extension services, and cognitive skills. *World Development 12,* 67-86.

Kalton, G., Lyberg, L., & Rempp, J. M. (1998). Review of methodology (Appendix A). In T. S. Murray, I. S. Kirsch, & L. B. Jenkins (Eds.), *Adult literacy in OECD countries: Technical report on the first International Adult Literacy Survey.* Report NCES 98-053. Washington, DC: U.S. Department of Education.

Kirsch, I., Jungeblut, A., Jenkins, L., & Kolstad, A. (1993). *Adult literacy in America: A first look at the results of the National Adult Literacy Survey.* Washington, DC: U.S. Department of Education, National Center for Education Statistics.

Kolstad, A. (1997). Information trade-offs in large scale assessments. In A. Tuijnman, I. Kirsch, & D. A. Wagner (Eds.), *Adult basic skills: Innovations in measurement and policy analysis.* Cresskill, NJ: Hampton Press.

Lavy, V., Spratt, J., & Leboucher, N. (1995). *Changing patterns of illiteracy in Morocco: Assessment methods compared.* LSMS Paper 115. Washington, DC: The World Bank.

Levine, K. (1998). Definitional and methodological problems in the cross-national measurement of adult literacy: The case of the IALS. *Written language and literacy* (Vol. 1 , pp. 41-61).

Little, A., & Wolf, A. (1996). *Assessment in transition: Learning monitoring and selection in international perspective.* London: Pergamon.

Murray, T. S. (1997). Proxy measurement of adult basic skills: Lessons from Canada. In A. Tuijnman, I. Kirsch, & D. A. Wagner (Eds.), *Adult basic skills: Innovations in measurement and policy analysis.* Cresskill, NJ: Hampton Press.

Murray, T. S., Kirsch, I. S., & Jenkins, L. B. (Eds.). (1998). *Adult literacy in OECD countries: Technical report on the first International Adult Literacy Survey* (Report NCES 98-053). Washington, DC: U.S. Department of Education.

OECD/Statistics Canada. (1995). *Literacy, economy and society.* Paris: OECD.

OECD/Statistics Canada. (1997). *Literacy skills for the knowledge society: Further results from the International Adult Literacy Survey.* Paris: OECD.

NCERT. (1995). *Minimum levels of learning at primary stage* (MLL). New Delhi: Government of India/NCERT.

Pike, K. L. (1966). *Language in relation to a unified theory of the structure of human behavior.* The Hague: Mouton.

Psacharopoulos, G., & Woodhall, M. (1985). *Education for development: An analysis of investment choices.* New York: Oxford University Press.

Reder, S. (in press). What does the NALS measure? Issues in dimensionality and construct validity. In M. C. Smith (Ed.), *Literacy for the 21st century: Research, policy and practice.* Boulder: Greenwood.

Scribner, S., & Cole, M. (1981). *The psychology of literacy.* Cambridge, MA: Harvard University Press.

Stigler, J. W., Schweder, R. A., & Herdt, G. (Eds.). (1990). *Cultural psychology.* New York: Cambridge University Press.

Tuijnman, A., Kirsch, I., & Wagner, D. A. (Eds.). (1997). *Adult basic skills: Innovations in measurement and policy analysis.* Cresskill, NJ: Hampton Press.

UNESCO. (1978). *Revised recommendation concerning the international standardization of educational statistics.* Paris: UNESCO.

UNESCO. (1983). *Statistics on educational attainment and literacy.* Paris: UNESCO.

UNESCO. (1990). *Final Report on the World Conference on Education For All: Meeting basic learning needs, Jomtien, Thailand.* Paris: UNESCO.

UNESCO. (1996a). *Data collection and analysis: Improving the information base for literacy programs. Basic Manual.* Paris: UNESCO.

UNESCO. (1996b). *Education For All: Mid-decade report.* Paris: UNESCO.

UNESCO/ILI. (1998). *Final report of expert seminar on out-of-school youth and adults.* Paris: Author.

Unicef. (1992). *Assessment of basic competencies of children in Bangladesh: A status paper.* Dhaka: Author.

United Nations Statistical Office (UNSO). (principal authors Wagner, D. A., & Srivastava, A. B. L.). (1989). *Measuring literacy through household surveys.* (Doc. No. DP/UN/INT-88-X01/10E). New York: United Nations Statistical Office.

U.S. Department of Education. (1996). *Third International Math/Science Study.* Washington, DC: Author.

Venezky, R. L. (1992). *Matching literacy testing with social policy: What are the alternatives?* (Policy Brief No. PB92-01). Philadelphia: University of Pennsylvania, National Center on Adult Literacy.

Venezky, R. L., Wagner, D. A., & Ciliberti, B. S. (Eds.). (1990). *Toward defining literacy.* Newark, DE: International Reading Association.

Vrignau, P., & Bonara, D. (1998). *Evaluation de la literatie et comparaisons internationales: Assurer la mesure, respecter les contextes.* In UNESCO/ILI Final Report of Expert Seminar on Literacy Assessment in Out-of-School Youth and Adults. Paris: Author.

Wagner, D. A. (1990). Literacy assessment in the Third World: An overview and proposed schema for survey use. *Comparative Education Review 33* (1), 112-138.

Wagner, D. A. (1992). *Literacy: Developing the future.* UNESCO Yearbook of Education. Paris: UNESCO.

Wagner, D. A. (1993). *Literacy, culture, and development: Becoming literate in Morocco.* New York: Cambridge University Press.

Wagner, D. A. (1995). Literacy and development: Rationales, myths, innovations, and future directions. *International Journal of Educational Development 15,* 341-362.

Wagner, D. A. (1998). Literacy retention: Comparisons across age, time and culture. In H. Wellman & S. G. Paris (Eds.), *Global prospects for education: Development, culture and schooling.* Washington, DC: American Psychological Association.

Wagner, D. A., & Venezky, R. L. (1998, in press). Adult literacy: The next generation. *Educational Researcher.*

Ziegahn, L. (1992). National literacy assessments: Present practice and alternatives. *International Journal of Educational Development 12,* 223-232.

INTERNATIONAL LITERACY INSTITUTE

PENN Graduate School of Education